



## RECLAMAÇÕES DE CLIENTES COMO FONTE DE INOVAÇÕES A PARTIR DE UMA BASE DE HELP DESK UTILIZANDO DATA MINING - UM EXEMPLO DE APLICAÇÃO.

Dayana Carla de Macedo (UTFPR, Brasil) [dayanamacedo@yahoo.com.br](mailto:dayanamacedo@yahoo.com.br)

Simone Nasser Matos (UTFPR, Brasil) [snasser@utfpr.edu.br](mailto:snasser@utfpr.edu.br)

Dálcio Roberto dos Reis (Universidade Positivo, Brasil) [dalcio.reis@gmail.com](mailto:dalcio.reis@gmail.com)

Daniel Poletto Tesser (UTFPR, Brasil) [danieltesser@utfpr.edu.br](mailto:danieltesser@utfpr.edu.br)

Libório de Oliveira Júnior (UTFPR, Brasil) [libajunior@gmail.com](mailto:libajunior@gmail.com)

**Resumo:** Em um mercado de acelerada competição e de consumidores cada vez mais exigentes, traçar o perfil do consumidor para adotar melhores estratégias de negociação pode fazer toda a diferença à medida que aumenta a satisfação do cliente. A presente pesquisa teve como objetivo a identificação do perfil de clientes para cada tipo de reclamações identificadas em uma base de dados sobre Help Desk de uma empresa de ração animal. Utilizou-se uma das técnicas de Mineração de Dados, chamada de clustering por meio do algoritmo k-means. A ferramenta selecionada para o respectivo trabalho foi o WEKA. Como resultados foram obtidos 23 clusters, com as respectivas características dos clientes para cada reclamação. Em alguns casos foram gerados 2 ou mais clusters para o mesmo tipo de reclamação, indicando perfis diferentes de clientes para o mesmo tipo de reclamação. Com o resultado, foi possível obter conhecimento das características de cada consumidor para cada tipo de reclamação, permitindo que a organização inove e atenda-os de forma customizada.

**Palavras chave:** Mineração de dados, Perfil de Clientes, Help Desk.

**Abstract:** In a fast market competition and increasingly demanding consumers, profiling the consumer to adopt better trading strategies can make all the difference as you increase customer satisfaction. This study aimed to identify the customer profile for each type of complaints identified in a database of Help Desk a pet food company. We used one of the techniques of Data Mining, called clustering using the k-means algorithm. The selected tool into their work was WEKA. As results were obtained 23 clusters, with their characteristics for each customer complaint. In some cases originated two or more clusters for the same type of complaint, indicating different profiles of customers for the same type of complaint. With the result, it was possible to obtain knowledge of the characteristics of each consumer for each type of complaint, allowing the organization to innovate and meet them in a customized manner.

**Keywords:** Data Mining, Customer Profile, Help Desk.

## 1. Introdução

Atualmente o mercado possui consumidores cada vez mais exigentes e, obter conhecimento sobre estes consumidores para adotar melhores estratégias de inovação, pode fazer toda a diferença. Desta forma, as organizações devem se preocupar em entender quais são os desejos, necessidades e expectativas em relação aos seus produtos a fim de manter os clientes atuais e captar os potenciais.

Segundo Rodrigues e Fernandez (2006), a competitividade global está obrigando as organizações a sempre reavaliar suas estratégias de atuação, com o intuito de garantir melhores resultados, sejam eles financeiros ou de participação de mercado.

Quando a questão refere-se aos clientes, o foco é estabelecer relacionamentos baseados no aprendizado de suas necessidades e desejos, oferecendo produtos certos e mantendo essa relação ao longo do tempo para a obtenção da identificação com a marca e garantia da lealdade por parte dos clientes (GAVA et al; 2005).

Nesse sentido verifica-se necessária a obtenção de conhecimento sobre os consumidores para o atendimento estratégico, de acordo com cada perspectiva individual. Logo, a presente pesquisa teve como objetivo identificar, como exemplo, o perfil dos clientes de uma fábrica de ração animal, para cada ocorrência ou reclamação registrada no Sistema Help Desk. A técnica utilizada para o experimento é a clustering, com o uso do algoritmo K-means.

## 2. Inovação focada no relacionamento com o Cliente

Prahalad e Krishnan (2008) afirmam que a nova Era da inovação é a inovação focada no relacionamento com o cliente. Esta nova Era mostra que a chave para a criação de valor e para o crescimento de todos os negócios no futuro depende do acesso a uma rede global de recursos para criar experiências singulares, com e para os clientes.

Desta forma, as empresas precisam inovar, pois um produto ou serviço inovador é aquele que oferece ao consumidor uma experiência única que o satisfaz fazendo com que ele busque esta empresa novamente. Os consumidores desejam adquirir produtos e serviços com custo atraente, qualidade perceptível e indiscutível, atendimento adequado e diferenciação (CARVALHO, REIS e CAVALCANTE, 2011).

Segundo Churchill & Peter (2000) pensamentos, sentimentos, ações e as influências sofridas pelas organizações refletem no comportamento do consumidor. Oportunamente as organizações podem escolher pela segmentação de mercado, que é o processo de dividir um mercado em grupos de compradores potenciais com desejos, necessidades, percepções de valores ou comportamentos de compra semelhantes tornando-se importante saber quem é o público alvo.

No entanto, a identificação de quem é o consumidor, sua posição atual ou pretendida na escala profissional, seu estilo de vida, suas



referências para o consumo e o que espera com a aquisição de um produto ou serviço, entre outros, fazem parte de um levantamento complexo.

Segundo Gordon (1999) o Marketing de relacionamento é o processo de identificar e criar novos valores com clientes individuais e o compartilhar de seus benefícios durante uma vida toda de parceria. Sendo uma espécie de condição que a empresa deve seguir para que ela saiba o que o cliente quer, criando um entendimento na mente do cliente do que a empresa pode lhe oferecer, ou seja, benefícios. Porém, o valor destes benefícios decorre de experiências singulares e personalizadas dos consumidores (PRAHALAD & KRISHNAN, 2008).

Já na visão de Kloter e Armstrong (2003), o Marketing de relacionamento significa criar, manter e aprimorar fortes relacionamentos com os clientes e outros interessados. Há a necessidade de executar estratégias para a atração de novos clientes, criar e manter transações com eles, além de reter os já existentes com o intuito de relacionamentos lucrativos e duradouros.

Segundo Crescitelli (2003) com a evolução do mercado, é identificada uma mudança no comportamento dos clientes acarretando uma mudança na concepção tradicional de marketing para evitar que fique obsoleta.

De acordo com Kotler (2000, p. 68):

"são mais difíceis de agradar. São mais inteligentes, mais conscientes em relação aos preços, mais exigentes, perdoam menos e são abordados por mais concorrentes com ofertas iguais ou

melhores".

Deve-se dar atenção especial no que diz respeito à manutenção de clientes já conquistados, pois a perda de um consumidor é mais prejudicial do que a "não-conquista" de um novo. Um consumidor perdido vale, efetivamente, por dois: é um cliente a menos para a empresa e um cliente a mais para o concorrente (CRESCITELLI, 2003).

Para Kotler (2000):

"A aquisição de novos clientes pode custar até cinco vezes mais do que os custos envolvidos em satisfazer e reter clientes existentes. Ela requer o emprego de um grande esforço para induzir clientes satisfeitos a deixar de contratar seus fornecedores atuais."

O marketing de relacionamento é caracterizado como uma espécie de filosofia empresarial que tem como previsão a construção e manutenção de relacionamentos individuais com os clientes, visando o estabelecimento de relações de longo prazo. Atualmente em mercados corporativos, onde o foco é o consumidor final, a sua adoção exige da organização o entrosamento por meio de uma postura cultural, estratégica e operacional que é calcada nos seus princípios (ANGELO, SCHNEIDER e LARÁN, 2006).

Para Mello e Leão (2006), o cliente associa valores ao produto ou serviço, ou seja, refere se a uma espécie de relação de custo/benefício, que consiste na diferença do que o cliente ganha adquirindo determinado produto menos o custo relacionado ao dinheiro, esforço, custo físico ou psíquico.



Para Ikeda e Oliveira (2005) quando uma organização visa compreender o significado do valor na oferta de um produto ou serviço, sob a perspectiva do cliente, há necessidade da criação e o aprimoramento de estratégias de marketing. Nesse sentido, torna-se notório compreender o que é o valor para o público alvo, para calcar as estratégias da organização. Dessa forma há a necessidade de estratégias bem sucedidas que demandam informações e conhecimento sobre seu público alvo.

Um serviço considerado bem sucedido pelo cliente necessita que a organização olhe para os produtos, empresa e métodos de serviço de acordo com os olhos do cliente (STERNE, 2000).

Uma vez obtido o modelo ou perfil do cliente, há possibilidade de utilizar essa informação para classificar novos clientes e consequentemente prever suas necessidades especiais (ROMDHANE, FADHEL & AYEB, 2010).

### **3. Mineração de dados**

Um dos grandes desafios de uma organização é manter uma carteira de clientes lucrativos. Com o conhecimento adquirido sobre os clientes é possível identificar seus objetivos, expectativas, perspectivas e desejos. Mas, para obter essa identificação, as ferramentas genéricas existentes no mercado, como Customer Relationship Management (CRM), Enterprise Resources Planning (ERP) ou Business Intelligence (BI) não são suficientes. Há necessidade da utilização de um método ou ferramenta com foco somente

na questão de obtenção de conhecimento dos consumidores, com capacidade analítica de identificar de padrões e prever, a partir dos dados estratégicos de uma organização, o comportamento de seus clientes, para o atendimento de forma customizada.

Segundo Braga (2005) isto é alcançado pela “mineração de dados (Data Mining) que é uma coleção de técnicas e métodos que facilitam a aquisição e retenção de informações da parte do mercado que a empresa possui interesse”.

A mineração de dados visa a extração de conhecimento de dados usando várias técnicas. Há várias aplicações bem sucedidas de mineração de dados em várias áreas incluindo WEB, Marketing, financeira e bancária, telecomunicação, entre outras. Com a aplicação da mineração as organizações podem desenvolver relacionamentos de longo prazo e bem sucedidos com seus clientes se há possibilidade de prever mudanças em seus hábitos de consumo (ou comportamento). No passado os pesquisadores normalmente utilizavam estatística para o estudo do comportamento de clientes, porém recentemente técnicas de mineração têm sido adotadas (ROMDHANE, FADHEL & AYEB, 2010).

As técnicas aplicam algoritmos específicos para extrair padrões com o intuito de obter informações que estão implícitas na base de dados, que são previamente desconhecidas e potencialmente úteis (CHEN, HEN & YU, 1996; MITRA, PAL & MITRA, 2002).

Dessa forma a mineração de dados utilizada para obtenção do perfil do cliente é uma tecnologia que permite construir modelos, descrevendo os



hábitos específicos, necessidades e o comportamento de um grupo de clientes. Uma vez obtido esse modelo, há possibilidade de serem usados para classificar novos clientes e, conseqüentemente, prever suas necessidades especiais (ROMDHANE, FADHEL & AYEB, 2010).

#### 4. Tarefa de clustering

Agrupamento é uma tarefa da mineração de dados que tem como objetivo segmentar uma população homogênea em um número de subgrupos ou de conjuntos mais homogêneos. Em um agrupamento não há nenhuma classe predefinida e nenhuma análise de exemplos. Os registros são agrupados em uma base de similaridade, neste caso, o usuário determina qual o significado, se houver algum, para a existência de cada cluster. Por exemplo, clusters de atributos de cliente podem indicar segmentos de mercado diferentes (SANTOS et al., 2007).

De acordo com Theodoridis (2003) para a tarefa de agrupamento há alguns passos que devem ser seguidos:

- Seleção de característica: essa fase tem como objetivo a identificação do subconjunto mais eficiente de características originais para se usar no processo de agrupamento.
- Medida de semelhança: refere-se à medida que vai quantificar a semelhança. É importante assegurar que todas as características contribuem igualmente para o cálculo da média de

semelhança e que não há característica que dominem outras.

- Critério de agrupamento: de acordo com o cluster que o analista deseja obter, este altera o critério para satisfazer suas expectativas.
- Algoritmos de agrupamento: com a medida de semelhança e os critérios de agrupamento definidos, determinar o algoritmo específico que irá montar a estrutura de agrupamento do conjunto de dados deve ser definido.
- Validação dos resultados: após os resultados obtidos pelo algoritmo de agrupamento é necessário verificar se aqueles dados estão corretos, sendo feito por meio de testes.
- Interpretação dos resultados: para chegar à conclusão às vezes o analista deve integrar aos resultados de agrupamento com outras evidências experimentais para se chegar às conclusões.

Santos et al. (2007) afirmam que é possível utilizar a aplicação de clustering para classificar clientes. Assim sendo, pode-se definir uma função que tem como parâmetros renda, idade, sexo, estado civil entre outros dados disponíveis na base. Porém, criar somente dois grupos, onde os clientes poderiam ter apenas os perfis A e B, pode não dizer muito sobre os consumidores, sendo importante encontrar o número correto de agrupamentos.

Dessa maneira a empresa deve evitar esse tipo de reclamação para essa parcela de consumidores



do seu mercado. Com o resultado é possível obter conhecimento das características de cada consumidor para cada tipo de reclamação, permitindo que a organização atenda-os de forma customizada. Após a definição de agrupamentos é necessário executar algum algoritmo de clustering sobre a base de dados.

#### 4.1 Algoritmos de Agrupamento

Na implementação desta tarefa são utilizados algoritmos tais como K-Means, K-Modes, K-Prototypes, K-Medoids, Kohonen (CARLANTONIO, 2001) dentre outros. Porém, para a identificação de perfil de clientes segundo Braga (2005), os mais recomendados são o K-Means e C-Means.

##### 4.1.1 ALGORITMO K-MEANS

O algoritmo K-means toma um parâmetro de entrada,  $k$ , e particiona um conjunto de  $n$  objetos em  $k$  cluster tal que a similaridade intracluster resultante é alta, porém, a similaridade intercluster é baixa. A similaridade de clusters refere-se à medida em relação ao valor médio dos objetos em um cluster, que pode ser visto como o centro de gravidade do cluster (CARLANTONIO, 2001).

Segundo Guerra (2006), o algoritmo K-médias (K-means clustering algorithm) tem um processamento simples de estimação dos parâmetros. Nesse sentido, suponha-se que existam  $N$  amostras (dados)  $x_n$  no total e se tem como objetivo encontrar  $K$  vetores  $c_j$  onde  $j = 1, \dots, K$ . O algoritmo procura particionar as amostras  $x_n$  em  $K$  subconjuntos  $S_j$  contendo  $N_j$  amostras, de tal maneira a minimizar a função da soma dos quadrados dados segundo Bishop (1995), conforme fórmulas a seguir.

$$J = \sum_{j=1}^K \sum_{x^n \in S_j} \|x^n - c_j\|^2 \quad (1)$$

Fonte: Bishop, (1995).

Onde  $c_j$  é o centro das amostras do conjunto  $S_j$  e é dado por:

$$c_j = \frac{1}{N_j} \sum_{x^n \in S_j} x^n \quad (2)$$

Fonte: Bishop, (1995).

O Quadro 1 resume o procedimento do algoritmo K-means segundo Carlantonio (2001).



## QUADRO 1 – ALGORITMO K-MEANS

**Algoritmo k-means** – o algoritmo k-means para particionamento baseia-se no valor médio dos objetos no cluster.

**Entrada:** O número de *clusters*,  $k$ , e a base de dados com  $n$  objetos.

**Saída:** Um conjunto de  $k$  *clusters* que minimizam o critério do erro quadrado.

**Método:**

1. Escolha arbitrariamente  $k$  objetos da base de dados como os centros iniciais dos *clusters*;
2. Repita
3. (Re) atribua cada objeto ao cluster ao qual o objeto é mais similar, de acordo com o valor médio dos objetos no cluster;
4. Atualize as médias dos *clusters*, isto é, calcule o valor médio dos objetos para cada cluster;
5. Até que não haja mudança de objetos de um cluster para outro.

Fonte: Carlan Antonio, (2001).

Segundo Carlan Antonio (2001) o algoritmo K-means faz a tentativa de determinação de  $k$  partições que minimizam a função do erro quadrado. Quando os clusters são densos e compactos e vem separados uns dos outros o algoritmo trabalha bem. Esse método é relativamente escalável e funciona de forma eficiente no processamento de grandes conjuntos de dados porque a complexidade computacional do algoritmo é  $O(n.k.t)$ , onde  $n$  é o número total de objetos,  $k$  é o número de clusters e  $t$  é o número total de interações.

O resultado deste método pode, em muitos casos, ser drasticamente afetado de acordo com a escolha das condições iniciais (KAINULAINEN, 2002). Porém, de acordo com Vale (2005), em bases de dados que estão bem estruturadas espera-se a convergência para um mínimo global. Quando o comportamento da convergência é lenta e os resultados de agrupamentos são bastante diferentes para diferentes configurações iniciais, isso pode indicar que o número de agrupamentos escolhido esteja errado, ou que os dados não possuam

estrutura de agrupamentos.

Segundo Costa (1999) o método apresentará bons resultados apenas quando os agrupamentos são hipersféricos e possuem aproximadamente o mesmo número de padrões em cada agrupamento. O desempenho do algoritmo dependerá também da escolha adequada da medida de distância e do ponto inicial de partida do algoritmo (VALE, 2005).

## 5. Help Desk ou SAC

O sistema Help Desk é assim explicado por Coêlho et al. (2003, p.46):

"[...] um sistema Help Desk constitui um mecanismo computacional facilitador de informação do tipo Help Desk = <Pergunta, Resposta> onde tanto a Pergunta em apreço se refere àquela de clientes e/ou usuários quaisquer; quanto a Resposta do sistema se refere a um apoio informacional bem definido em seu domínio".



Dois fatores estão presentes: automatização da criação de chamados para concessão lógica de acesso ainda dependentes de ação manual e automatização de criação de chamados de instalação de pré-requisito para acessos.

Os Help Desk são usados atualmente para qualquer tipo de assistência via telefone, seguindo uma dinâmica específica: os operadores recebem o chamado e interpretam o problema. Com base em sua experiência e conhecimento auxiliam o cliente prestando informações ou recomendando ações que resolvam o problema (SILVA, 2007).

Segundo Silva (2007), um sistema de Help Desk é composto por três componentes básicos:

"Software: auxilia o controle do inventário tecnológico da empresa, revelando os problemas repetidos, tempos médios de atendimento, as soluções usadas nas áreas mais demandantes e identificação da necessidade de treinamentos.

Equipe: atua nos níveis de campo e de retaguarda. A equipe de campo faz o atendimento aos problemas, solucionando-os quando possível ou acionando fornecedores internos e externos. Já a equipe de retaguarda é acionada quando um problema técnico supera a capacidade de resolução da equipe de campo. Os analistas de suporte de fornecedores funcionam como uma terceira camada e podem ser requisitados pela equipe de retaguarda, recebendo o problema identificado e já mapeado.

Metodologia de serviço: envolve a

situação do sistema de Help Desk em relação aos usuários, definindo quais são as estratégias de ação diante de determinado problema, apresentando indicadores de desempenho e identificando novas oportunidades em um processo de melhoria contínua".

Coêlho et al. (2003) associam o Help Desk à tecnologia de gestão do conhecimento, enfatizando a sua modelagem com o emprego de metodologias e ferramentas variadas.

O Help Desk tem como objetivo principal evitar a perda de tempo do usuário ao se deparar com alguma dificuldade com a tecnologia. A rápida e eficiente intervenção favorece a redução do impacto no negócio, como propostas não enviadas, sistema bancário inativo e demais consequências que surgem. É característica dos departamentos de Help Desk a busca de novos métodos para o aumento de eficiência, em especial, o de promover ao usuário a autoajuda, localizando soluções e orientações por si mesmo (COHEN, 2008).

## 6. Metodologia

Com o intuito de identificar o perfil dos consumidores de acordo com cada tipo de reclamação, utilizou-se uma das técnicas da mineração de dados para efetuar a clustering das reclamações de acordo com o perfil dos clientes. Com relação a classificação metodológica a presente pesquisa quanto à natureza é aplicada, pois visa gerar conhecimento para aplicação prática de solução de específicos problemas. Já





do ponto de vista dos objetivos é Explicativa, pois de acordo com Gil (1991), em geral as pesquisas que são experimentais assumem a forma de explicativa, implicando a identificação dos fatores que determinam ou contribuem para a ocorrência dos fenômenos. No aspecto dos procedimentos técnicos caracteriza-se como uma pesquisa experimental e bibliográfica, pelo fato do levantamento teórico para fundamentação da pesquisa e experimento devido à determinação do objeto de estudo, bem como, seleção de variáveis para observação dos efeitos que a variável produz. Para finalizar em relação à abordagem é quantitativa, pois para Gil (1991) para essa classificação a pesquisa traduz números e/ou opiniões em informação para classificá-las e analisá-las.

Para a realização do experimento seria necessária a obtenção da base de dados da organização, contendo os dados dos clientes desta. Porém, não se teve acesso a esses dados. Dessa forma, para a realização do experimento, foi selecionada uma base de dados de domínio público, que contém os dados e os atributos de pessoas adultas, semelhante à base de dados da organização para a realização do experimento. Esta base de dados está disponível para download a partir do site Machine Learning Repository. A ferramenta selecionada para o respectivo trabalho foi o WEKA, que é uma ferramenta gratuita, desenvolvida pela Universidade de Waikato.

Com o intuito de categorizar essas pessoas de acordo com as reclamações efetuaram-se o acréscimo das possíveis reclamações de forma aleatória para o uso experimental a essa base. O arquivo total possui 32579 instâncias.

A preparação da base de dados foi dividida em três tarefas: aquisição de uma base de dados contendo dados de pessoas adultas, levantamento das principais ocorrências que são registradas em sistemas de Help Desk em uma fábrica de ração animal e por último, a construção de uma base única contendo os dados fictícios de perfis de clientes e as ocorrências que foram feitas em um sistema de Help Desk. Para a tarefa de montar uma única base para a mineração dos dados, foi utilizado um programa de computador o qual adicionou aleatoriamente à base de clientes um tipo de ocorrência ou reclamação para cada cliente.

## 6. Análise e Resultados

A busca dos tipos de reclamações, iniciou-se com uma pesquisa bibliográfica em uma indústria de ração animal para cães e gatos (linha pet) de médio porte, localizada no sudoeste do Paraná, buscando, em documentos de Instrução de Trabalho, registros de não-conformidades com os tipos de ocorrências utilizados por esta indústria. Nesta pesquisa identificaram-se os seguintes tipos de ocorrências: Aglomeração de finos com gorduras; Aglomeração de grãos; Contaminação na produção por mistura de produtos; Cor; Corpo estranho; Corpo estranho insetos; Corpo estranho metálico; Corpo estranho vidros; Data de validade; Densidade; Excesso de oleosidade; Excesso de pó; Formato irregular; Informação sobre formulação; Intoxicação; Mofo; Óbito; Odor anormal; Problemas de pelagem; Reações cutâneas;



Tamanho; Umidade e Vômito.

A ferramenta selecionada para o respectivo trabalho foi o WEKA, que é uma ferramenta gratuita, desenvolvida pela Universidade de Waikato. O Quadro 2 ilustra a descrição dos atributos.

**QUADRO 2: DESCRIÇÃO DOS ATRIBUTOS**

Atributos	Descrição
Idade	Real
Classe de Trabalho	Privado, Autônomo, Público Federal, Público Municipal, Público Estadual, Não possui trabalho, Nunca trabalhou.
Peso final	Real.
Educação	Bacharel, Ensino Superior, 11th <sup>1</sup> , Nível de Ensino Médio, Escola Profissional, 9th <sup>2</sup> , 7th <sup>3</sup> -8th <sup>4</sup> , Mestrado, 1st <sup>5</sup> -4th <sup>6</sup> , 10th <sup>7</sup> , Doutor (a), 5th <sup>8</sup> - 6th <sup>9</sup> , Pré-escola.
Num - educação	Real.
Estado Civil	Casado com Cônjuge Civil, Divorciado, Solteiro, Separado, Viúvo, Casado com Cônjuge Ausente, Casado com Cônjuge

<sup>1</sup> 11th refere-se ao décimo primeiro ano *High School* nos Estados Unidos, equivalente ao Ensino Médio no Brasil;

<sup>2</sup> 9th refere-se ao nono ano do *High School* nos Estados Unidos, equivalente ao Ensino Médio no Brasil;

<sup>3</sup> 7th refere-se ao sétimo ano do *Kindergarten* nos Estados Unidos, equivalente ao Jardim de Infância no Brasil;

<sup>4</sup> 8th refere-se ao oitavo ano do *Kindergarten* nos Estados Unidos, equivalente ao Jardim de Infância no Brasil;

<sup>5</sup> 1st refere-se ao primeiro ano do *Kindergarten* nos Estados Unidos, equivalente ao Jardim de Infância no Brasil

<sup>6</sup> 4th refere-se ao quarto ano do *Kindergarten* nos Estados Unidos, equivalente ao Jardim de Infância no Brasil

<sup>7</sup> 10th refere-se ao décimo ano do *High School* nos Estados Unidos, equivalente ao Ensino Médio no Brasil;

<sup>8</sup> 5th refere-se ao quinto ano do *Kindergarten* nos Estados Unidos, equivalente ao Jardim de Infância no Brasil

<sup>9</sup> 6th refere-se ao sexto ano do *Kindergarten* nos Estados Unidos, equivalente ao Jardim de Infância no Brasil

	das forças armadas.
Ocupação	Suporte Técnico, Artesanato, Outro Serviço, Vendas, Executivo-gerencial, Professor especialista, Doméstica, Inspetor de Máquinas, Padre, Agricultura-pesca, Transporte, Serviços privados - casa, Serviços de Segurança, Forças Armadas.
Relação	Esposa, Enteado ou Enteada, Esposo, Não possui vínculo familiar, Outro vínculo familiar, Solteiro.
Raça	Branco, Habitante de Ilha, Esquimó, Negro.
Sexo	Feminino, Masculino.
Ganho de capital	Real.
Perda de Capital	Real.
Horas por semana de Trabalho	Real.
País Nativo	Estados Unidos, Camboja, Inglaterra, Porto Rico, Canadá, Alemanha, Periféricos EUA, Índia, Japão, Grécia, Sul, China, Cuba, Irã, Honduras, Filipinas, Itália, Polônia, Jamaica, Vietnã, México, Portugal, Irlanda, França, Republica Dominicana, Laos, Equador, Taiwan, Haiti, Columbia, Hungria, Guatemala, Nicarágua, Escócia, Tailândia, Iugoslávia, El-Salvador, Trinidad e Tobago, Peru, Hong-Kong, Holanda.
Classe	Óbito, Vômito, Intoxicação, Aglomeração de grãos, Aglomeração de Finos <sup>10</sup> com Gorduras, Cor, Densidade, Tamanho, Excesso de oleosidade,

<sup>10</sup> Finos: Pode ser entendido como a porção da ração peletizada, ou seja, parte da ração que se encontra desagregada de sua estrutura inicial, resultando em partículas de dimensões menores que os peletes (KLEIN, 1996).



	Excesso de pó, Formato Irregular, Informação sobre Formulação, Contaminação na produção por mistura de produtos, Mofo, Corpo estranho, insetos, Corpo Estranho metálico, Corpo Estranho Vidros, Odor anormal, Problemas de pelagem. Data de validade, Reações cutâneas, Umidade.
--	--

Fonte: Os autores (2011)

Para o experimento foi criado um arquivo chamado `perfil_sac.arff`, pois a ferramenta WEKA trabalha com arquivos nessa extensão. O arquivo contém a seguinte estrutura:

- Relação: Sendo a primeira linha do arquivo, onde deve conter a variável `@relation` seguida de uma palavra-chave que tem como intuito identificar a relação ou tarefa que está sendo executada.
- Atributos: é o conjunto de linhas, onde cada linha é iniciada com `@attribute` seguindo do nome do atributo e seu tipo.
- Dados: Posteriormente a uma linha contendo `@data`. Onde cada linha corresponde a uma instância, devendo ter valores separados por vírgula correspondentes dos atributos seção `@attribute`.

O algoritmo utilizado foi o K-means, sendo chamado na ferramenta WEKA de Simplekmeans. Para o entendimento do funcionamento do algoritmo K-means, a base de dados possui uma tabela com colunas e linhas, contendo os dados os quais serão categorizados. Na tabela a coluna é denominada de dimensão, ou seja, cada coluna refere-se a um atributo e

cada linha contém informações sobre esses atributos. O algoritmo K-means analisou todos os dados dessas tabelas e criou as classificações, onde este o algoritmo Simplekmeans indicou um cluster com as respectivas linhas pertencentes a este.

Porém, para a geração dos clusters o algoritmo K-means fez uma comparação entre cada valor de cada linha por meio da distância, neste caso, distância euclidiana <sup>11</sup>. Posteriormente a esse cálculo das distâncias, algoritmo efetuou o cálculo dos centróides para cada um dos clusters. De acordo com Pichiliani (2006), um centróide é calculado pela média dos valores de cada atributo de cada ocorrência que é pertencente a esta centróide. Onde o algoritmo gera k centróides e coloca as ocorrências da tabela de acordo com sua distância dos centróides.

Como resultados foram gerados 23 clusters, sendo que cada cluster correspondente a cada tipo de reclamação. Para a visualização dos dados segue abaixo uma legenda para os atributos e para as reclamações (Tabela 1). Vale lembrar que cada cluster gerado por reclamação definirá uma série de informações sobre o cliente (atributos).

A Tabela 1 contém os clusters gerados para a

---

<sup>11</sup> Distância euclidiana é usada como uma medida de distância, ou seja, medidas de separabilidade de divergência ou discriminação. Desta forma, para duas classes, um atributo  $X_i$  é preferido ao atributo  $X_j$  se  $X_i$  provê uma diferença maior que  $X_j$  entre as probabilidades condicionais das duas classes. Na medida euclidiana, ocorre à comparação entre cada valor de cada linha por meio da distância euclidiana que mede o quanto longe uma ocorrência está da outra (LEE, 2005).



reclamação ‘Odores anormais’.

**TABELA 1: DESCRIÇÃO DE ATRIBUTOS**

Tributos	Reclamações
A1: Idade	R1: Odor anormal
A2: Classe de Trabalho	R2: Óbito
A3: Peso Final	R3: Excesso de pó
A4: Educação	R4: Tamanho
A5: Num - Educação	R5: Excesso de oleosidade
A6: Estado Civil	R6: Densidade
A7: Ocupação	R7: Corpo estranho
A8: Relação	R8: Aglomeração de finos com gorduras
A9: Raça	R9: Contaminação na produção por mistura de produtos
A10: Sexo	R10: Corpo estranho metálico
A11: Ganho de Capital	R11: Corpo estranho insetos
A12: Perda de Capital	R12: Reações cutâneas
A13: Horas por semana	R13: Mofo
A14: País Nativo	R14: Cor
	R15: Vômito
	R16: Formato irregular
	R17: Aglomeração de grãos
	R18: Corpo estranho vidros;
	R19: Data de validade;
	R20: Informação sobre formulação;
	R21: Intoxicação;
	R22: Problemas de pelagem;
	R23: Umidade

Fonte: Os Autores (2011).

**TABELA 2: RECLAMAÇÃO ODOR ANORMAL**

Class	R1	R1	R1	R1
A1	38,58	40,95	43,46	47,63
A2	Privado	Autônomo	Privado	Autônomo
A3	189.778,37	159.901,28	186.799,24	182.264,33
A4	Nível de Ensino Médio	Nível de Ensino Médio	Bacharel	Mestre
A5	10,08	9,07	13,24	13,53
A6	Casado-civ-Conjugê	Nunca casou	Casado-civ-Conjugê	Divorciado
A7	Prof-Especialista	Agricultura e Pesca	Prof-Especialista	Prof-Especialista
A8	Marido	Não-em-família	Marido	Não-em-família
A9	Branco	Branco	Branco	Branco
A10	M	M	M	M
A11	1.077,65	1.362,66	4.134,16	2.870,22
A12	87,3	75,58	17,3	86,97
A13	40,44	50,63	45,68	43,68
A14	EUA	EUA	EUA	EUA

Fonte: Os Autores (2011).

Conforme se pode verificar na Tabela 2 foram gerados 4 clusters para a reclamação Odor Anormal, indicando que, para esse tipo de reclamação, há 4 perfis diferentes de clientes. Ou seja, os clientes que reclamaram do produto com relação a odor anormal possuem diferenças entre si, acarretando a geração de 4 clusters. Pelo fato de não terem similaridades entre si, não foi possível gerar somente 1 cluster, resultando no

agrupamento de similaridades de informações dos clientes em 4 clusters. Para esse tipo de reclamação podemos notar que os clientes são do sexo masculino, possuem um alto nível de escolaridade, onde dos 3 clusters gerados os consumidores atuam na área de docência. A seguir seguem as Tabelas 3 e 4 contendo os dois clusters gerados para a reclamação Óbito e Excesso de pó:

**TABELA 3: RECLAMAÇÃO ÓBITO**

Class	R2	R2
A1	44,36	44,88
A2	Gov-Local	Privado
A3	170.456,86	196.538,66
A4	Mestre	Alguma - Universidade
A5	13,2	9,68

**TABELA 4: RECLAMAÇÃO EXCESSO DE PÓ**

Class	R3	R3
A1	45,45	25,73
A2	Privado	Privado
A3	191.753,00	205.697,03
A4	10th	Alguma - Faculdade
A5	6,71	9,81

A6	Casado-civ-Conjugê	Divorciado
A7	Prof-Especialista	Prof-Especialista
A8	Esposa	Solteiro
A9	Branco	Preto
A10	F	F
A11	2.321,23	358,11
A12	166,66	36,02
A13	40,43	36,23
A14	EUA	EUA

Fonte: Os autores (2011).

A6	Casado-civ-Conjugê	Nunca casou
A7	Transporte	Padre
A8	Esposo	Própria Criança
A9	Branco	Preto
A10	M	F
A11	698,77	259,95
A12	94,97	48,42
A13	43,8	34,01
A14	EUA	EUA

Fonte: Os autores (2011).

De acordo com os dados na Tabela 3 foram gerados 2 clusters para a reclamação óbito, sendo essa a reclamação mais crítica de todas, demonstrando que a organização deve atentar para o perfil do cliente dessa reclamação. Também nota-se que para essa reclamação os consumidores são todos do sexo feminino e

possuem um alto grau de escolaridade e atuam na docência. Para a reclamação de excesso de pó, os consumidores possuem diferenças opostas em relação a sexo e área de atuação. Já nas tabelas 5 e 6 seguem os clusters gerados para a reclamação de tamanho e excesso de oleosidade.

TABELA 5: RECLAMAÇÃO TAMANHO

Class	R4	R4
A1	46,76	36,79
A2	Privado	Privado
A3	188.655,63	189.184,64
A4	Alguma - Universidade	Nível de Ensino Médio
A5	9,8	9,56
A6	Divorciado	Divorciado
A7	Vendas	Padre
A8	Não-em-familia	Solteiro
A9	Branco	Branco
A10	M	F
A11	735,75	290,99
A12	95,14	40,08
A13	43,73	40,1

TABELA 6: RECLAMAÇÃO EXCESSO DE OLEOSIDADE

Class	R5	R5
A1	32,17	56,21
A2	Privado	Privado
A3	186.138,66	176.786,33
A4	Nível de Ensino Médio	Nível de Ensino Médio
A5	8,49	9,27
A6	Nunca Casou	Casado-civ-Conjugê
A7	Outro-Serviço	Prof-Especialista
A8	Solteiro	Esposo
A9	Branco	Branco
A10	F	M
A11	189,11	1.497,69
A12	28,55	68,52
A13	34,4	38,19



A14 EUA EUA

Fonte: Os autores (2011).

A14 EUA EUA

Fonte: Os autores (2011).

Para ambas as reclamações percebe-se que os consumidores atuam no setor privado, porém não possuem alto grau de escolaridade. Pois, para a reclamação excesso de oleosidade verifica-se que esses possuem somente ensino médio e para a reclamação tamanho foram gerados somente 2 clusters, porém, somente em um deles o

consumidor apresenta o nível médio, demonstrando, dessa forma, o baixo nível de escolaridade.

Na tabela 7 podemos verificar os 2 clusters gerados para a reclamação aglomeração de finos com gorduras:

**TABELA 7: AGLOMERAÇÃO DE FINOS COM GORDURAS**

Class	R8	R8
A1	34,6	48,91
A2	Privado	Autônomo
A3	188.640,08	181.546,54
A4	Bacharel	Nível de Ensino Médio
A5	12,36	8,41
A6	Nunca casou	Separado
A7	Prof-Especialista	Outro Serviço
A8	Não-em-Família	Solteiro
A9	Branco	Branco
A10	F	F
A11	767,99	147,6
A12	69,64	65,02
A13	39,46	35,66
A14	EUA	EUA

Fonte: Os autores (2011).

Por meio dos dados identificou-se que foram gerados dois clusters para a reclamação de finos com gorduras, resultando em 2 perfis de clientes para esse tipo de reclamação. Ambos possuem gênero feminino, porém apresentam diferenças com relação ao grau de escolaridade. Abaixo seguem as Tabelas 8 e 9 contendo os clusters gerados para as reclamações, onde somente obteve um 1 cluster por reclamação. Por meio dos dados verificamos que os perfis entre os clientes diferem entre si com relação ao atributo idade, estado civil, grau de escolaridade, ocupação, porém apresentam algumas similaridades tais como o gênero masculino, país nativo, horas

trabalhado durante a semana e renda. Já nos clusters gerados na Tabela 8 verificamos que para as reclamações de cor, formato irregular e aglomeração de grãos os resultados apontam o mesmo grau de escolaridade para essas reclamações, todos atuam no setor privado. Para as reclamações de cor, formato irregular e aglomeração de grãos os consumidores possuem somente o nível médio e atuam em áreas distintas, como artesanato, inspeção de máquinas e serviços de limpeza. Ou seja, a organização, teria ou deveria atentar que para esse público evitando tais reclamações, pela questão da não reincidência da mesma reclamação.

**TABELA 8: CLUSTERS GERADOS PARA CADA RECLAMAÇÃO**

Class	R12	R13	R14	R16	R17
A1	39,85	30,72	35,59	42,6	29,36
A2	Privado	Privado	Privado	Privado	Privado
A3	181.375,72	194.223,45	188.393,33	181.511,05	211.742,98
A4	Alguma - Universidade	Alguma - Universidade	Nível de Ensino Médio	Nível de Ensino Médio	Nível de Ensino Médio
A5	9,93	10,65	9,1	9,13	8,53
A6	Casado-civ- Conjugê	Nunca casou	Casado-civ- Conjugê	Casado-civ- Conjugê	Nunca casou
A7	Executivo Gerencial	Vendas	Inspetor de Maquinas	Artesanato	Doméstica
A8	Esposa	Não em família	Esposo	Esposo	Não em família
A9	Branco	Branco	Branco	Branco	Branco
A10	F	M	M	M	M
A11	1.085,41	670,16	656,23	784,97	198,32
A12	100,49	65,05	47,67	108,29	52,88
A13	36,9	41,81	44,9	43,91	41,41
A14	EUA	EUA	EUA	EUA	EUA

Fonte: Os autores (2011).



Através dos resultados verificou-se que, para as reclamações de vômito, corpo estranho vidros, data de validade, informação sobre formulação, intoxicação, problema de pelagem e umidade não foram gerados clusters. Porém essas reclamações foram identificadas na base de dados, mas não foram gerados clusters, por que as informações ficaram dispersas entre si, não agrupadas o suficiente a ponto de gerar um cluster. Ainda, salienta-se que foram gerados mais de um cluster para a mesma reclamação, o que demonstra que essas reclamações foram mais impactantes que as que não foram geradas clusters.

De todas as reclamações a mais impactante foi a Odor anormal com 4 clusters gerados, seguido de óbito, excesso de pó, tamanho, excesso de oleosidade e aglomeração de finos, todos com 2 clusters gerados.

Nesse sentido através dos resultados obtidos, se os dados fossem reais de uma organização, esta poderia verificar que deve focar esforços imediatos na redução dessas reclamações, principalmente com relação aqueles que foram gerados maiores números de clusters. Ainda por meio dos resultados, a organização teria conhecimento do perfil do cliente de cada reclamação e atentar para evitar novas reclamações com esses clientes. Pois atualmente em um mercado altamente competitivo a reincidência de uma falha, por exemplo, como o óbito do animal, pela ingestão de um produto que essa empresa comercializa, acarreta a perda desse e de potenciais clientes.

De acordo com a conclusão da execução do

algoritmo k-means o responsável pela gestão passa a conhecer as características de cada consumidor, ou seja, o perfil do cliente para cada reclamação de forma rápida e simples. Como exemplo, através dos resultados pode-se perceber os clientes que possuem titulação de mestrado, possuem reclamação para óbito e odor anormal. Dessa maneira é possível verificar que clientes com alto grau de escolaridade reclamaram do odor anormal da ração e de óbito. Dessa forma a empresa deve evitar esse tipo de reclamação para essa parcela de consumidores do seu mercado. Com o resultado foi possível obter conhecimento das características de cada consumidor para cada tipo de reclamação, permitindo que a organização atenda-os de forma customizada. Dessa forma, o gestor passa a ter o conhecimento que deve evitar determinados tipos de reclamações para determinados tipos de clientes. Com isto pode melhor oferecer os serviços e produtos aos seus clientes.

## **8. Considerações Finais**

Para obter conhecimento das características e comportamento de clientes, através da mineração de dados, usou-se a tarefa de clustering. A tarefa de categorização “Clustering” tem como objetivo segmentar uma população homogênea em um número de subgrupos ou de conjuntos mais homogêneos. Dos algoritmos da tarefa de agrupamento os algoritmos K-means e C-means são os mais utilizados, porém para a presente pesquisa foi utilizado o algoritmo K-means, pois, segundo a



literatura, para esse tipo de tarefa é o mais recomendado.

Para a aplicação da tarefa de clustering foi selecionada uma base de dados de domínio público, que está disponível para download a partir do site Machine Learning Repository contendo dados de pessoas adultas. Com o intuito de categorizar essas pessoas de acordo com as reclamações de uma empresa de ração de animais, onde foi efetuado o acréscimo das possíveis reclamações de forma aleatória para o uso experimental da base de dados.

Após a execução do algoritmo K-means como resultado obteve-se 23 clusters com as respectivas características dos clientes para cada reclamação. Assim a empresa deve evitar esse tipo de reclamação para essa parcela de consumidores do seu mercado. Com o resultado foi possível obter conhecimento das características de cada consumidor para cada tipo de reclamação, permitindo que a organização atenda-os de forma customizada. Foram gerados 2 ou mais clusters para o mesmo tipo de reclamação, indicando perfis diferentes de clientes para o mesmo tipo de reclamação. Ou seja, pelos dados não foi possível agrupar as similaridades dos clientes em somente 1 cluster, resultando em mais clusters.

Desta forma, o gestor pode verificar, para determinado tipo de cliente, qual reclamação deve ser evitada. Observou-se que a utilização da mineração de dados em conjunto com um sistema Help Desk é viável, com vantagens no que diz respeito ao atendimento customizado, visto que é possível categorizar os clientes de

acordo com as respectivas reclamações, e, dessa maneira, o gestor passa a ter acesso a ter o conhecimento de quais reclamações deve evitar para determinados tipos de clientes, possibilitando um atendimento estratégico.

Como possíveis dificuldades apontam-se a questão da análise dos resultados, ou seja, a análise da base de dados, com o intuito da compreensão do significado e relevância dos atributos. Também se salienta a importância de levantar inicialmente um objetivo, o que se deseja obter com a aplicação da ferramenta, qual conhecimento deseja-se obter. Exigindo-se um entendimento da situação atual do ambiente em que será realizado o processo. É necessário que haja interação entre os gestores e área de TI, ou seja, o analista de TI.

Para a organização que tem em mente realizar aplicação desse tipo de ferramenta, deve inicialmente efetuar um levantamento inicial dos objetivos que almeja alcançar com o processo. Porém é necessário que esta detenha um profissional de TI para efetuar essa tarefa juntamente com os gestores, para que ocorra a interatividade desses profissionais melhorando os resultados obtidos. Bem como a definição dos objetivos, planejamento da execução das atividades e análise dos resultados.

### Referências

- BISHOP, Christopher. M. (1995). **Neural Networks for Pattern Recognition**. Clarendon Press. Oxford.
- BRAGA, Luis. P. V. B. **Introdução à Mineração de Dados**. 2ª Edição Revista e Ampliada. Rio de Janeiro: E-Papers Serviços Editoriais, 2005.



- CARLANTONIO, Lando Mendonça di. **Novas metodologias para clusterização de dados.** Dissertação (Mestrado)-Engenharia Civil, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001. Disponível em: <  
[http://wwwp.coc.ufjf.br/teses/mestrado/inter/2002/teses/di%20CARLANTONIO\\_LM\\_02\\_t\\_M\\_int.pdf](http://wwwp.coc.ufjf.br/teses/mestrado/inter/2002/teses/di%20CARLANTONIO_LM_02_t_M_int.pdf) > acesso em: 10 fev. 2010.
- CARVALHO, H.G.; REIS, D.R.;CAVALCANTE, M.B. **Gestão da Inovação.** Curitiba: Editora Aymará, 2011.
- CHEN, M.S.; HEN, J.; YU, P. S. (1996). **Data Mining: An overview from a database perspective.** JEEE Transactions on knowledge Data Engineering, 8 (6), 866-883Disponível em: <  
[http://cs.nju.edu.cn/zhoush/zhoush.files/course/dm/reading/reading01/chen\\_tkde96.pdf](http://cs.nju.edu.cn/zhoush/zhoush.files/course/dm/reading/reading01/chen_tkde96.pdf) >
- CHURCHILL, Gilberto A.; PETER, J. Paul. **Marketing Criando Valor para os Clientes.** São Paulo: Saraiva, 2000.
- COÊLHO, Álvaro Vinícius de Souza; FERNEDA, Edilson; MARTINS, Agenor de Sousa; BARROS, M. A.; GORGONIO, F. L. **Help Desk inteligente em gestão do conhecimento: um tratamento integrador de paradigmas.** Inesc Em Revista, Unaí, MG, v. 1, n. 1, p. 46-51, 2003.
- COHEN, Roberto. **Implantação de help desk e service desk.** São Paulo: Novatec Editora, 2008.
- COSTA, José Alfredo Ferreira. **Classificação Automática e Análise de Dados por Redes Neurais Auto-Organizáveis.** Campinas, SP, 1999. Disponível em: <  
<http://cutter.unicamp.br/document/?code=vtls000195524> > acesso em 20 mai. 2010.
- CRESCITELLI, Edson. **Marketing de relacionamento: Um estudo sobre as relações entre vendedores varejistas e fabricantes de eletrodomésticos.** 2003. 213 f. Tese (Doutorado). Faculdade de economia, Administração e Contabilidade, Departamento de Administração, Universidade de São Paulo, São Paulo, 2003. Disponível em: <  
<http://www.teses.usp.br/teses/disponiveis/12/12139/tde-27042009-175454/pt-br.php> > acesso em 19 mai. 2010.
- D'ANGELO, André Cauduro; SCHNEIDER, Heleno; LARÁN, Julianio A. **Marketing de Relacionamento junto a Consumidores Finais: Um estudo exploratório com grandes empresas brasileiras.** Revista de Administração Contemporânea, RAC. V.10; n.1; Jan./ Mar. 2006, p. 73-93. Disponível em: <  
<http://www.scielo.br/pdf/rac/v10n1/a05.pdf> > acesso de 14 mai. 2010.
- GAVA, V. L.; SPINOLA, M. M.; NOMURA, L.; GONÇALVES, R. F. **Comércio eletrônico: aspectos que devem ser considerados em sua análise/ implementação e avaliação no mercado brasileiro.** In: Encontro Nacional de Engenharia de Produção, 2005, Poro Alegre. XXV ENEGEP, 2005. Disponível em: <  
[http://www.abepro.org.br/biblioteca/ENEGE P2005\\_Enegep0904\\_0787.pdf](http://www.abepro.org.br/biblioteca/ENEGE P2005_Enegep0904_0787.pdf) > acesso em: 10 abr. 2010.
- GIL, Antonio Carlos. **Como elaborar projetos de pesquisa.** São Paulo: Atlas, 1991.
- GOMES, H. G.; REIS, D. R.; CAVALCANTE, M. B. **Gestão da Inovação.** Curitiba: Aymará, 2011.
- GORDON, Ian. **Marketing de Relacionamento: Estratégias, Técnicas e Tecnologias para Conquistar Clientes e Mantê-los para Sempre.** São Paulo. Editora Futura, 1999.
- IKEDA, A. A.; OLIVEIRA, T. M. V. de. **O Conceito de valor para o cliente: Definições e implicações gerenciais em marketing.** REAd - Edição 44 Vol. 11 nº. 2, mar-abr 2005.
- KAINULAINEN, Jukka. **Clustering Algorithms: basics and Visualition,** 2002. Disponível em: <  
<http://www.niksula.cs.hut.fi/~jkainula/pdfs/clustering.pdf> >. Acesso em 14 dez. 2009.
- KLEIN, C.H. **Efeito da forma física e do nível de energia da ração sobre o desempenho, a composição de carcaça e a eficiência de utilização da energia metabolizável consumida por frangos de corte.** 1996. 97f. Dissertação (Mestrado). Universidade Federal do Rio Grande do Sul.



- KOTLER, Philip. **Administração de Marketing**. 10. Ed. São Paulo: Pearson Prentice Hall, 2000. Disponível em: <<http://www.fae.edu/publicacoes/pdf/mkt/1.pdf>> acesso em: 10 abr. 2010.
- KOTLER, Philip; ARMSTRONG, Gary. **Princípios de Marketing**. Tradução Arlete Simile Marques, Sabrina Cairo, 9ª edição. São Paulo. Prentice Hall, 2003. P. 204 - 206.
- LEE, S.; PARK, Y. **Customization of technology roadmaps according to roadmapping purposes**: Overall process and detailed modules. *Technological Forecasting & Social Change*, v. 72, p. 567-583, 2005. Disponível em: <[http://www.maoner.com/Cited\\_Saritas\\_One\\_r\\_2004.pdf](http://www.maoner.com/Cited_Saritas_One_r_2004.pdf)> Acesso em: 5 jul. 2011.
- MELLO, Sérgio Carvalho Benício de ; LEÃO, André Luiz Maranhão de Souza. **Uma Revisão Crítica sobre o Uso de Diferentes Conceitos de Valor no Marketing**. *RAC Eletrônica (Online)*, v. 2, p. 37-56, 2008. Disponível em: <<http://www.slideworld.net/online-journal.aspx/Uma-Revis%C3%A3o-Cr%C3%ADtica-sobre-o-Uso-de-Diferentes-Conceitos-de-Valor-no-Marketing-ppt-13069>> acesso em 24 abr. 2010.
- MITRA, S.; PAL, S. K.; MITRA, P. **Data mining in soft computing Framework: a survey**. *IEEE Transactions On Neural Networks*, v. 13, n. 1, p. 3-14, 2002. Disponível em: <<http://repository.ias.ac.in/26054/1/310.pdf>>. Acesso em: 6 ago. 2011.
- PICHILIANI, M. **Data Mining na Prática: Algoritmo K-means**. 2006. Disponível em: <<http://www.devmedia.com.br/post-4584-Data-Mining-na-Pratica-Algoritmo-K-Means.html>> acesso em 29 de novembro de 2010.
- PRAHALAD, C. K.; KRISHNAN, M. S. **A nova era da Inovação: A inovação focada no relacionamento com o cliente**. Rio de Janeiro: Elsevier, 2008.
- RODRIGUES, L. C.; FERNANDEZ, M. J. **Alinhamento Estratégico da Tecnologia de Informação e Inteligência Competitiva**. *Revista Gestão Industrial (Online)*, Ponta Grossa - PR, v. 2, n. 1, p. 131-146, 2006. Disponível em: <<http://www.pg.cfetpr.br/depog/periodicos/index.php/revistagi/article/viewArticle/135>> acesso em 13 nov. 2010.
- ROMDHANE, L. B.; FADHEL, N.; AYE, B. (2010). **An efficient approach for building customer profiles from business data**. *Expert Systems with Applications*, 37, 1573-1585. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0957417409006125>>. Acesso em: 13 nov. 2010.
- SANTOS, Alfredo. **Gerenciamento de identidade**. Rio de Janeiro: Brasport, 2007. Disponível em: <[http://books.google.com.br/books?id=ACFU300zVGUC&printsec=copyright&source=gbs\\_pub\\_info\\_s&cad=2#v=onepage&q=&f=false](http://books.google.com.br/books?id=ACFU300zVGUC&printsec=copyright&source=gbs_pub_info_s&cad=2#v=onepage&q=&f=false)>. Acesso em: 27 ago. 2009.
- SANTOS, Fernando, A, S.; CLEMENTE, Fernando, O.; SOBROZA, Luciano, A.; YOKOYAMA, Miriam, S. **Data Mining aplicado á Análise de carteira de Ações**. Monografia (Graduação) - Escola politécnica da Universidade de São Paulo. São Paulo, São Paulo, 2007. Disponível em: <[http://www.pcs.usp.br/~pcspf/2007/Semestral%202007/PCS%202502%20SEM%20Grupo%20\(07\)/grupo7s.pdf](http://www.pcs.usp.br/~pcspf/2007/Semestral%202007/PCS%202502%20SEM%20Grupo%20(07)/grupo7s.pdf)> acesso em: 25 nov. 2010.
- SILVA, Jorge Moacir Farias da. **Utilização do raciocínio baseado em casos como apoio a um sistema de Help Desk**. 2007. Monografia (Graduação em Sistemas de Informação). Centro Universitário Feevale - Instituto de Ciências Exatas e Tecnológicas - Novo Hamburgo, 2007. Disponível em: <<http://nead.feevale.br/tc/files/1001.doc>>. Acesso em: 27 ago. 2009.
- STERNE, J. **Customer service on the internet: building relationships, increasing loyalty, and staying competitive**. 2nd ed. New York: John Wiley, 2000.
- THEODORIDIS, Sergios.; KOUTROUMBAS, Konstantinos. **Pattern recognition**. Elsevier Academic Press, 2003.
- VALE, Marcos, N. do. **Agrupamentos de Dados: Avaliação de Métodos e Desenvolvimento de Aplicativo para Análise de Grupos**. Dissertação (Mestrado) - Pós Graduação em Engenharia Elétrica da Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro, Rio de Janeiro, 2005. Disponível em: <[http://www.maxwell.lambda.ele.puc-rio.br/Busca\\_etds.php?strSecao=resultado&nrSeq=7975@](http://www.maxwell.lambda.ele.puc-rio.br/Busca_etds.php?strSecao=resultado&nrSeq=7975@)> acesso em: 25 nov. 2009.